(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 110413201 A
(43)申请公布日 2019.11.05

(21)申请号 201810399624.3

(22)申请日 2018.04.28

(71)申请人 伊姆西IP控股有限责任公司
地址 美国马萨诸塞州

(72)发明人 刘友生 徐鑫磊 杨利锋 高健 李雄成

(74)专利代理机构 北京市金杜律师事务所 11256
代理人 王茂华 李峥宇

(51)Int.Cl.
G06F 3/06(2006.01)
G06F 12/0866(2016.01)

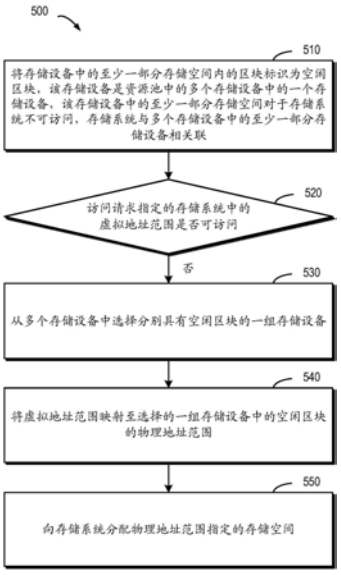
权利要求书3页 说明书12页 附图10页

(54)发明名称

用于管理存储系统的方法、设备和计算机程序产品

(57)摘要

本公开涉及一种用于管理存储系统的方法、设备和计算机程序产品。提供了一种用于管理存储系统的方法，存储系统与来自资源池中的多个存储设备中的至少一部分存储设备相关联，多个存储设备中的存储设备中的至少一部分存储空间对于存储系统不可访问。该方法包括：将存储设备中的至少一部分存储空间内的区块标识为空闲区块；响应于确定访问请求指定的存储系统中的虚拟地址范围不可访问，从多个存储设备中选择分别具有空闲区块的一组存储设备；将虚拟地址范围映射至选择的一组存储设备中的空闲区块的物理地址范围；以及向存储系统分配物理地址范围指定的存储空间。由此，以更为方便并且快捷的方式实现资源池的扩展，进而提高存储系统的性能。



CN 110413201 A

1.一种用于管理存储系统的方法,所述存储系统与来自资源池中的多个存储设备中的至少一部分存储设备相关联,所述多个存储设备中的存储设备中的至少一部分存储空间对于所述存储系统不可访问,所述方法包括:

将所述存储设备中的所述至少一部分存储空间内的区块标识为空闲区块;

响应于确定访问请求指定的所述存储系统中的虚拟地址范围不可访问,从所述多个存储设备中选择分别具有空闲区块的一组存储设备;

将所述虚拟地址范围映射至选择的所述一组存储设备中的所述空闲区块的物理地址范围;以及

向所述存储系统分配所述物理地址范围指定的存储空间。

2.根据权利要求1所述的方法,其中所述存储系统是基于独立磁盘冗余阵列的存储系统,其中从所述多个存储设备中选择分别具有空闲区块的一组存储设备包括:

基于所述独立磁盘冗余阵列的配置来确定所述一组存储设备的数量;以及

基于所述数量来选择所述一组存储设备。

3.根据权利要求1所述的方法,其中基于所述数量来选择所述一组存储设备包括:

从所述多个存储设备中确定包括空闲区块的存储设备的空闲设备数量;

响应于确定所述空闲设备数量满足所述数量,从所述多个存储设备中选择包括空闲区块的存储设备以作为所述一组存储设备。

4.根据权利要求3所述的方法,其中基于所述数量来选择所述一组存储设备进一步包括:

响应于确定所述空闲设备数量不满足所述数量,将所述存储设备以外第一存储设备中的第一区块中的数据移动至所述存储设备;

将所述第一区块标识为空闲区块;以及

增加所述空闲设备数量。

5.根据权利要求1所述的方法,进一步包括:

将所述存储设备以外的第一存储设备中的第一区块中的数据移动至所述存储设备;以及

将所述第一区块标识为空闲区块。

6.根据权利要求1所述的方法,其中所述存储系统是基于独立磁盘冗余阵列的存储系统,所述存储系统包括多个切片,所述方法进一步包括:

响应于接收到针对所述多个切片中的切片的切片访问请求,

基于所述存储系统的切片分配表确定所述切片的地址范围;

基于所述切片的所述地址范围确定所述虚拟地址范围;以及其中向所述存储系统分配所述物理地址范围指定的存储空间包括:从分配的所述存储空间中选择存储空间以用于所述切片。

7.根据权利要求6所述的方法,其中基于所述切片的所述地址范围确定所述虚拟地址范围包括:

确定与所述切片的所述地址范围相关联的位图的地址,所述位图中的相应位指示所述多个切片中的相应切片中的数据是否为零;以及

基于所述位图的所述地址来确定所述虚拟地址范围。

8.根据权利要求7所述的方法,进一步包括:

在分配的所述存储空间中选择位图空间以用于存储所述位图;以及

响应于确定所述用户访问的为写操作,

向所述切片写入所述由写操作指定的目标数据;以及

将所述位图中的与所述切片相关联的位设置为指示所述切片中的数据为非零。

9.根据权利要求8所述的方法,其中所述存储设备是在所述资源池的扩展期间被插入所述资源池的新存储设备;以及

所述新存储设备中的至少一部分存储空间的物理地址尚未与所述存储系统建立地址映射关系。

10.根据权利要求9所述的方法,进一步包括:

针对所述多个切片中的每个切片生成切片访问请求。

11.一种用于管理存储系统的设备,包括:

至少一个处理器;

易失性存储器;以及

与所述至少一个处理器耦合的存储器,所述存储器具有存储于其中的指令,所述指令在被所述至少一个处理器执行时使得所述设备执行用于管理存储系统的动作,所述存储系统与来自资源池中的多个存储设备中的至少一部分存储设备相关联,所述多个存储设备中的存储设备中的至少一部分存储空间对于所述存储系统不可访问,所述动作包括:

将所述存储设备中的所述至少一部分存储空间内的区块标识为空闲区块;

响应于确定访问请求指定的所述存储系统中的虚拟地址范围不可访问,从所述多个存储设备中选择分别具有空闲区块的一组存储设备;

将所述虚拟地址范围映射至选择的所述一组存储设备中的所述空闲区块的物理地址范围;以及

向所述存储系统分配所述物理地址范围指定的存储空间。

12.根据权利要求11所述的设备,其中所述存储系统是基于独立磁盘冗余阵列的存储系统,其中从所述多个存储设备中选择分别具有空闲区块的一组存储设备包括:

基于所述独立磁盘冗余阵列的配置来确定所述一组存储设备的数量;以及

基于所述数量来选择所述一组存储设备。

13.根据权利要求11所述的设备,其中基于所述数量来选择所述一组存储设备包括:

从所述多个存储设备中确定包括空闲区块的存储设备的空闲设备数量;

响应于确定所述空闲设备数量满足所述数量,从所述多个存储设备中选择包括空闲区块的存储设备以作为所述一组存储设备。

14.根据权利要求13所述的设备,其中基于所述数量来选择所述一组存储设备进一步包括:

响应于确定所述空闲设备数量不满足所述数量,将所述存储设备以外第一存储设备中的第一区块中的数据移动至所述存储设备;

将所述第一区块标识为空闲区块;以及

增加所述空闲设备数量。

15.根据权利要求11所述的设备,其中所述动作进一步包括:

将所述存储设备以外的第一存储设备中的第一区块中的数据移动至所述存储设备;以及

将所述第一区块标识为空闲区块。

16.根据权利要求11所述的设备,其中所述存储系统是基于独立磁盘冗余阵列的存储系统,所述存储系统包括多个切片,所述方法进一步包括:

响应于接收到针对所述多个切片中的切片的切片访问请求,

基于所述存储系统的切片分配表确定所述切片的地址范围;

基于所述切片的所述地址范围确定所述虚拟地址范围;以及

其中向所述存储系统分配所述物理地址范围指定的存储空间包括:从分配的所述存储空间中选择存储空间以用于所述切片。

17.根据权利要求16所述的设备,其中基于所述切片的所述地址范围确定所述虚拟地址范围包括:

确定与所述切片的所述地址范围相关联的位图的地址,所述位图中的相应位指示所述多个切片中的相应切片中的数据是否为零;以及

基于所述位图的所述地址来确定所述虚拟地址范围。

18.根据权利要求17所述的设备,其中所述动作进一步包括:

在分配的所述存储空间中选择位图空间以用于存储所述位图;以及

响应于确定所述用户访问的为写操作,

向所述切片写入所述由写操作指定的目标数据;以及

将所述位图中的与所述切片相关联的位设置为指示所述切片中的数据为非零。

19.根据权利要求18所述的设备,其中所述存储设备是在所述资源池的扩展期间被插入所述资源池的新存储设备;以及

所述新存储设备中的至少一部分存储空间的物理地址尚未与所述存储系统建立地址映射关系。

20.根据权利要求19所述的设备,其中所述动作进一步包括:

针对所述多个切片中的每个切片生成切片访问请求。

21.一种计算机程序产品,所述计算机程序产品被有形地存储在非瞬态计算机可读介质上并且包括机器可执行指令,所述机器可执行指令用于执行根据权利要求1-10中的任一项所述的方法。

用于管理存储系统的方法、设备和计算机程序产品

技术领域

[0001] 本公开的各实现方式涉及存储管理，更具体地，涉及用于管理存储系统(例如，独立磁盘冗余阵列，Redundant Array of Independent Disks，RAID)的方法、设备和计算机程序产品。

背景技术

[0002] 随着数据存储技术的发展，各种数据存储设备已经能够向用户提供越来越高的数据存储能力，并且数据访问速度也有了很大程度的提高。在提高数据存储能力的同时，用户对于数据可靠性和存储系统的响应时间也提供了越来越高的需求。目前，已经开发出了基于冗余磁盘阵列系统的多种数据存储系统来提高数据的可靠性。当存储系统中的一个或者多个磁盘出现故障时，可以从其他正常操作的磁盘上的数据来恢复出故障磁盘中的数据。

[0003] 目前已经开发出了映射独立磁盘冗余阵列(Mapped RAID)。在该映射RAID中，磁盘是一个逻辑概念并且可以包括多个区块(extent)。一个逻辑磁盘中包括的多个区块可以分布在资源池中的不同物理存储设备上。对于映射RAID的一个条带中的多个区块而言，该多个区块应当分布在不同的物理存储设备上，以便当该多个区块中的一个区块所在的物理存储设备出现故障时，可以执行重建操作以便从其他区块所在的物理存储设备中恢复数据。

[0004] 在存储系统的使用期间，用户可能会从资源池中请求分配更多的存储空间，并且新用户还可能从资源池中请求分配存储空间以创建新的存储系统。这些资源分配请求可能会导致资源池中的空闲空间逐渐降低甚至可能被耗尽。此时，需要向资源池中加入新的存储设备，以便扩充资源池的存储空间。

[0005] 需要基于新的存储设备中的各个区块的地址来更新存储系统的地址映射，以使得新的存储设备中的存储空间对于存储系统的用户可用。然而，需要基于新存储设备中的各个区块的地址来逐步更新地址映射，这将造成用户需要等待地址映射更新完毕才能使用资源池中新增加的存储空间。因而，如何以更为方便并且快捷的方式实现资源池的扩展，进而提高存储系统的性能，成为一个技术难题。

发明内容

[0006] 因而，期望能够开发并实现一种以更为有效的方式来管理存储系统的技术方案。期望该技术方案能够与现有的存储系统相兼容，并且通过改造现有存储系统的各种配置，来以更为有效的方式管理存储资源池。

[0007] 根据本公开的第一方面，提供了一种用于管理存储系统的方法，存储系统与来自资源池中的多个存储设备中的至少一部分存储设备相关联，多个存储设备中的存储设备中的至少一部分存储空间对于存储系统不可访问。该方法包括：将存储设备中的至少一部分存储空间内的区块标识为空闲区块；响应于确定访问请求指定的存储系统中的虚拟地址范围不可访问，从多个存储设备中选择分别具有空闲区块的一组存储设备；将虚拟地址范围映射至选择的一组存储设备中的空闲区块的物理地址范围；以及向存储系统分配物理地址

范围指定的存储空间。

[0008] 根据本公开的第二方面，提供了一种用于管理存储系统的设备，包括：至少一个处理器；易失性存储器；以及与至少一个处理器耦合的存储器，存储器具有存储于其中的指令，指令在被至少一个处理器执行时使得设备执行用于管理存储系统的动作。存储系统与来自资源池中的多个存储设备中的至少一部分存储设备相关联，多个存储设备中的存储设备中的至少一部分存储空间对于存储系统不可访问。该动作包括：将存储设备中的至少一部分存储空间内的区块标识为空闲区块；响应于确定访问请求指定的存储系统中的虚拟地址范围不可访问，从多个存储设备中选择分别具有空闲区块的一组存储设备；将虚拟地址范围映射至选择的一组存储设备中的空闲区块的物理地址范围；以及向存储系统分配物理地址范围指定的存储空间。

[0009] 根据本公开的第三方面，提供了一种计算机程序产品，计算机程序产品被有形地存储在非瞬态计算机可读介质上并且包括机器可执行指令，机器可执行指令用于执行根据本公开的第一方面的方法。

附图说明

[0010] 结合附图并参考以下详细说明，本公开各实现方式的特征、优点及其他方面将变得更加明显，在此以示例性而非限制性的方式示出了本公开的若干实现方式。在附图中：

[0011] 图1A和1B分别示意性示出了其中可以实现本公开的方法的存储系统的示意图；

[0012] 图2示意性示出了其中可以实现本公开的方法的示例性环境的框图；

[0013] 图3示意性示出了图2中的存储资源池的图示；

[0014] 图4示意性示出了根据本公开的一个实现方式的用于管理存储系统的架构图；

[0015] 图5示意性示出了根据本公开的一个实现方式的用于管理存储系统的方法的流程图；

[0016] 图6示意性示出了根据本公开的一个实现方式的用于从多个存储设备中选择用于分配给存储系统的区块的框图；

[0017] 图7示意性示出了根据本公开的一个实现方式的用于从多个存储设备向存储系统分配存储空间的方法的流程图；

[0018] 图8A和图8B示意性示出了根据本公开的一个示例性实现的用于在多个存储设备之间移动数据的框图；

[0019] 图9A示意性示出了根据本公开的一个示例性实现的存储系统的结构的示意图；

[0020] 图9B示意性示出了根据本公开的一个示例性实现的存储系统中的切片与位图之间的关系的示意图；

[0021] 图10示意性示出了根据本公开的一个示例性实现的用于管理存储系统的设备的框图；以及

[0022] 图11示意性示出了根据本公开的一个示例性实现的用于管理存储系统的设备的框图。

具体实施方式

[0023] 下面将参照附图更详细地描述本公开的优选实现。虽然附图中显示了本公开的优

选实现，然而应该理解，可以以各种形式实现本公开而不应被这里阐述的实现所限制。相反，提供这些实现是为了使本公开更加透彻和完整，并且能够将本公开的范围完整地传达给本领域的技术人员。

[0024] 在本文中使用的术语“包括”及其变形表示开放性包括，即“包括但不限于”。除非特别申明，术语“或”表示“和/或”。术语“基于”表示“至少部分地基于”。术语“一个示例实现”和“一个实现”表示“至少一个示例实现”。术语“另一实现”表示“至少一个另外的实现”。术语“第一”、“第二”等等可以指代不同的或相同的对象。下文还可能包括其他明确的和隐含的定义。

[0025] 在本公开的上下文中，存储系统可以是基于RAID的存储系统。基于RAID的存储系统可以将多个存储设备组合起来，成为一个磁盘阵列。通过提供冗余的存储设备，可以使得整个磁盘组的可靠性大大超过单一的存储设备。RAID可以提供优于单一的存储设备的各种优势，例如，增强数据整合度，增强容错功能，增加吞吐量或容量，等等。RAID存在多个标准，例如RAID-1，RAID-2，RAID-3，RAID-4，RAID-5，RAID-6，RAID-10，RAID-50等等。关于RAID级别的更多细节，本领域技术人员例如可以参见https://en.wikipedia.org/wiki/Standard_RAID_levels、以及https://en.wikipedia.org/wiki/Nested_RAID_levels等。

[0026] 图1A示意性示出了其中可以实现本公开的方法的存储系统100A的示意图。在图1A所示的存储系统中，以包括五个独立存储设备(110、112、114、116以及118)的RAID-5(4D+1P，其中4D表示存储系统中包括四个存储设备来用于存储数据，1P表示存储系统中包括一个存储设备来用于存储P校验)阵列为示例，来说明RAID的工作原理。应当注意，尽管图1A中示意性示出了五个存储设备，在其他的实现方式中，根据RAID的等级不同，还可以包括更多或者更少的存储设备。尽管图1A中示出了条带120、122、124、…、126，在其他的示例中，RAID系统还可以包括不同数量的条带。

[0027] 在RAID中，条带跨越多个物理存储设备(例如，条带120跨越存储设备110、112、114、116以及118)。可以简单地将条带理解为多个存储设备中的满足一定地址范围的存储区域。在条带120中存储的数据包括多个部分：存储在存储设备110上的数据块D00、存储在存储设备112上的数据块D01、存储在存储设备114上的数据块D02、存储在存储设备116上的数据块D03、以及存储在存储设备118上的数据块P0。在此示例中，数据块D00、D01、D02、以及D03是被存储的数据，而数据块P0是被存储数据的P校验。

[0028] 在其他条带122和124中存储数据的方式也类似于条带120，不同之处在于，有关其他数据块的校验可以存储在不同于存储设备118的存储设备上。以此方式，当多个存储设备110、112、114、116以及118中的一个存储设备出现故障时，可以从其他的正常的存储设备中恢复出故障设备中的数据。

[0029] 图1B示意性示出了存储系统110A的重建过程的示意图100B。如图1B所示，当一个存储设备(例如，以阴影示出的存储设备116)出现故障时，可以从其余的正常操作的多个存储设备110、112、114、118中恢复数据。此时，可以向RAID中加入新的后备存储设备118B来替代存储设备118，以此方式，可以将恢复的数据写入118B并实现系统的重建。

[0030] 应当注意，尽管在上文中参见图1A和图1B描述了包括5个存储设备(其中4个存储设备用于存储数据，1个存储设备用于存储校验)的RAID-5的存储系统，根据其他RAID等级的定义，还可以存在包括其他数量的存储设备的存储系统。例如，基于RAID-6的定义，可以

利用两个存储设备来分别存储校验P和Q。又例如，基于三重校验RAID的定义，可以利用三个存储设备来分别存储校验P、Q和R。

[0031] 随着分布式存储技术的发展，图1A和1B所示的存储系统中的各个存储设备110、112、114、116以及118可以不再局限于物理存储设备，而是可以是虚拟存储设备。例如，存储设备110上的各个区块可以分别来自于资源池中的不同的物理存储设备(在下文中将简称为存储设备)。图2示意性示出了其中可以实现本公开的方法的示例性环境的框图。如图2所示，存储资源池270可以包括多个物理存储设备210、220、230、240、250、…、260。此时，该多个存储设备中的存储空间可以被分配给多个存储系统290、…、292。此时，存储系统290、…、292可以经由网络280来访问存储资源池270中的各个存储设备中的存储空间。

[0032] 图3示意性示出了如图2所示的存储资源池270的更多信息的图示。资源池270可以包括多个存储设备210、220、230、240、250、…、260。每个存储设备可以包括多个区块，例如，在每个存储设备上部示出了每个存储设备中包括的区块的示意图。空白区块(如图例360所示)表示空闲的区块，以条纹示出的区块(如图例362所示)表示用于图1中的存储系统110A的第一条带的区块，而以阴影示出的区块(如图例364所示)表示用于图1中的存储系统110A的第二条带的区块。此时，用于第一条带的区块312、322、332、342、352分别用于存储第一条带的数据块D00、D01、D02、D03和校验P0。用于第二条带的区块324、334、344、366和314分别用于存储第二条带的数据块D10、D11、D12、D13和校验P1。

[0033] 如图3所示，在各个存储设备中还可以存在预留的空闲部分370，以便用于在资源池中的一个存储设备出现故障时，可以选择各个存储设备中的空闲部分370中的区块，来重建故障存储设备中的各个区块。

[0034] 应当注意，图3仅以4D+1P的RAID-5存储系统为示例示出了各个条带中的区块如何分布在资源池的多个存储系统中。当采用基于其他RAID等级时，本领域技术人员可以基于上文的原理来实现具体细节。例如，在6D+1P+1Q的RAID-6存储系统中，每个条带中的8个区块可以分布在多个存储设备上，进而保证多个存储设备的负载均衡。

[0035] 将会理解，为了扩展资源池270中的存储空间，可以向该资源池270中加入新的存储设备。图4示意性示出了根据本公开的一个实现方式的用于管理存储系统的架构图400。图4示意性示出了存储系统290和292，该存储系统290和292分别具有地址映射294和296，以用于记录各个存储系统与资源池270中的存储设备之间的地址映射关系。

[0036] 假设期望建立新的存储系统420，当发现资源池270中的存储空间不足时，可以向资源池270中加入新的存储设备410，以便为存储系统420提供存储空间。需要借助于地址映射422来访问存储系统420。为了确保资源池270中的各个存储设备的负载平衡，可以将原有存储设备210、220、230、240、250、…、260中的数据移动至新的存储设备410中，进而使得被分配的存储空间在当前资源池270中以尽可能均匀的方式分布。可以将上述数据移动过程称为“混洗(shuffle)”操作。然而，由于存储系统与资源池270中的存储设备之间的地址映射尚未完成，因而导致用户并不能访问存储系统。

[0037] 为了解决上述缺陷，本公开的实现方式提供了一种用于管理存储系统420的方法、设备和计算机程序产品。具体地，根据本公开内容的一个实现方式，提供了一种用于管理存储系统420的方法，存储系统420与来自资源池270中的多个存储设备中的至少一部分存储设备相关联，多个存储设备中的存储设备410中的至少一部分存储空间对于存储系统420不

可访问。

[0038] 如图4所示，存储系统420与来自资源池270中的多个存储设备210、220、230、240、250、…、260以及410中的至少一部分存储设备相关联。如图4所示，存储设备410是在资源池270的扩展期间新加入的存储设备，并且该多个存储设备中的存储设备(例如，存储设备410)中的至少一部分存储空间对于存储系统420不可访问。此时，存储系统420与资源池270中的各个存储设备之间的地址映射422尚未完成建立，因而，按照传统的技术方案，用户并不能访问存储系统420中的存储空间，而是必须等到地址映射422被完全建立。

[0039] 在此实现中，在存储设备410被加入资源池270之后，可以将存储设备410中的至少一部分存储空间内的区块标识为空闲区块。此时，如果确定访问请求指定的存储系统420中的虚拟地址范围不可访问，可以从多个存储设备中选择分别具有空闲区块的一组存储设备。在此实现中，不必等待整个资源池270的混洗操作的完成，可以优先地在资源池270中寻找空闲空间来响应于访问请求。

[0040] 接着，可以将虚拟地址范围映射至选择的一组存储设备中的空闲区块的物理地址范围，以便完成地址映射422中的与访问请求相关联的部分的建立。继而，向存储系统420分配物理地址范围指定的存储空间。利用上述示例性实现，可以基于资源池270中的空闲空间，来完成地址映射422中的至少一部分，进而使得可以访问存储系统420中的与该部分相关的访问。以此方式，可以降低存储系统420的用户的等待时间，进而提高存储系统420的性能。将会理解，通常而言，当资源池270中出现存储空间的短缺时，将会向资源池270中加入一个或多个新的存储设备。利用上述示例性实现，可以尽量利用资源池270中的零散的存储空间，来服务于访问请求以降低等待时间。

[0041] 图5示意性示出了根据本公开的一个实现方式的用于管理存储系统的方法500的流程图。在框510处，可以将存储设备410中的至少一部分存储空间内的区块标识为空闲区块。将会理解，在执行本公开的方法期间，混洗操作可以并行地运行。因而存储设备410中的一部分存储空间可能已经由于混洗操作而完成了地址映射的更新。此时，存储设备410中已经经历了混洗操作的部分已经被映射至其他存储系统290、…、292，并且相应的地址映射294和296已经被更新。因而，存储设备410未被混洗操作占用的空间是未被分配的，可以将该部分存储空间内的区块标识为空闲区块。

[0042] 如果接收到针对存储系统420中的虚拟地址范围的访问请求，可以确定该虚拟地址范围是否可访问。如果不可访问，则该方法500前进至框530。在框530处，可以从多个存储设备210、220、230、240、250、…、260以及410中选择分别具有空闲区块的一组存储设备。继而，在框540处，可以将虚拟地址范围映射至选择的一组存储设备中的空闲区块的物理地址范围。进一步，在框550处，可以向存储系统420分配物理地址范围指定的存储空间。此时，由于此时已经针对存储系统420中的在框550处分配的存储空间建立了地址映射，因而该被分配的存储空间可以用户访问。换言之，在地址映射422被完全建立之前，存储系统420中的一部分存储空间可以对外界提供服务。

[0043] 根据本公开的一个示例性实现，存储系统420可以是基于RAID的存储系统。此时，可以基于独立磁盘冗余阵列的配置来确定一组存储设备的数量，并且基于数量来从资源池270中选择一组存储设备。将会理解，基于RAID的存储系统可以利用数据冗余来提供更高的可靠性。因而此时需要基于RAID的具体配置来确定从资源池270中选择来自多少个存储设

备的空闲区块来建立地址映射。

[0044] 继续上文的示例，将以4D+1P的存储系统为示例描述本公开的更多信息。当存储系统420为4D+1P的RAID时，可以从5个(4+1＝5)存储设备中分别选择一个空闲区块。在选择的5个空闲区块中，4个空闲区块用于存储数据，而1个空闲区块用于存储与数据相关联的校验。根据本公开的一个示例性实现，在基于6D+1P+1Q的存储系统中，可以从8个(6+1+1＝8)存储设备中选择空闲区块。在下文中，将参见图6描述如何建立地址映射的更多细节。

[0045] 图6示意性示出了根据本公开的一个实现方式的用于从多个存储设备中选择用于分配给存储系统的区块的框图600。该图6示出了存储设备410刚被加入资源池270的情况，此时，存储设备410中的全部区块都可以被标识为空闲区块。在图6中，空白区块(如图例670所示)表示空闲区块，网格区块(如图例672所示)表示已经被使用的区块。此时，可以分别从存储设备210、220、230、240、250、…、260以及410中选择包括空闲区块的存储设备。

[0046] 如图6所示，存储设备210、220、230、240和410分别包括空闲区块610、620、630、640和650，因而可以分别选择这些存储设备和相应的区块，以用于建立地址映射422。在此实现中，可以将空闲区块610、620、630、640和650的空间与虚拟地址范围660建立映射关系，进而将上述空闲区块分配给存储系统420。可以在地址映射422中记录虚拟地址范围660与各个空闲区块之间的映射关系。

[0047] 可以采用多种格式来描述地址映射422。例如，可以以每个空闲区块的全局唯一的标识符来指示该空闲区块。可以基于多种方式来构建全局唯一的标识符。根据本公开的一个示例性实现，对于区块610而言，可以以该区块610所在的存储设备210的编号以及该区块610在存储设备210中的位置来作为标识符。此时，区块610的标识符可以是ID＝(device:210;position:2)，表示该区块610是设备210中的第二个区块。可以以类似方式确定其他空闲区块的标识符。例如，区块620的标识符可以是ID＝(device:220;position:3)，表示该区块620是设备220中的第三个区块。根据本公开的一个示例性实现，还可以以全局唯一的方式来对各个存储设备中的区块设置标识符。

[0048] 由于存储系统420是基于4D+1P的RAID，此时可以在空闲区块610、620、630、640中存储数据，并且在空闲区块650中存储数据的校验。以此方式，一方面可以优先地为存储系统420中的一部分存储空间建立地址映射关系，另一方面还可以通过提供冗余存储的方式，确保存储系统420中的数据的可靠性。

[0049] 根据本公开的一个示例性实现，可以在资源池270中确定包括空闲区块的存储设备的空闲设备数量。如果确定空闲设备数量满足数量，从多个存储设备中选择包括空闲区块的存储设备以作为一组存储设备。对于存储系统420而言，如果确定资源池270中包括空闲区块的存储设备的数量大于或者等于5，则可以从资源池中选择5个包括空闲区块的存储设备，并且按照上文描述的过程来执行分配。

[0050] 根据本公开的一个示例性实现，如果确定空闲设备数量不满足数量，则可以将存储设备以外第一存储设备中的第一区块中的数据移动至存储设备410，并且将第一区块标识为空闲区块。此时，可以增加空闲设备数量，并且判断当前的空闲设备数量是否满足执行地址映射所需的区块的数量。如果当前空闲设备数量不能满足所需数量，则可以将包括空闲区块的存储设备以外的其他存储设备中的区块中的数据移动至存储设备410，以使得资源池270中包括存在包括空闲区块的更多存储设备。

[0051] 在下文中，将参见图7描述有关分配存储空间的更多细节。图7示意性示出了根据本公开的一个实现方式的用于从多个存储设备向存储系统420分配存储空间的方法700的流程图。如图7所示，在框710所示，可以基于RAID的配置来确定一组存储设备的数量。例如，对于4D+1P的RAID而言，一组存储设备的数量可以是4+1＝5；对于4D+1P+1Q的RAID而言，一组存储设备的数量可以是4+2＝6。

[0052] 在框720处，可以从多个存储设备中确定包括空闲区块的存储设备的空闲区块数量。在框730处，如果空闲区块的数量满足所需数量，则方法前进至框760，此时即可从多个存储设备中选择分别具有空闲区块的一组存储设备。如果空闲区块的数量不满足所需数量，则方法前进至框740。此时，可以在资源池270中执行混洗操作，以便使得资源池270中的更多存储设备中包括空闲区块。接着，在框750处，可以增加空闲设备数量。例如，在框730处确定的空闲设备数量为4，此时在框740处已经将一个存储设备中的区块中的数据移动至新增的存储设备410后，则空闲设备数量将变为4+1＝5。

[0053] 继而，方法返回值框730，并且在此可以判断空闲设备数量是否满足所需数量。由于此时空闲设备数量已经被更新至5并且等于所需数量，方法前进至760。继而，可以从多个存储设备中选择分别具有空闲区块的一组存储设备。

[0054] 在上文中已经描述在发现资源池270中的包括空闲区块的存储设备的数量不满足所需数量时执行数据混洗操作的实现。根据本公开的一个示例性实现，还可以与上文描述的方法500并行地执行数据混洗操作。换言之，可以并行地执行常规的混洗操作，以便从资源池270中的原有存储设备210、220、230、240、250、…、260中的数据移动至新增的存储设备410。通过混洗操作，资源池270中的更多存储设备将包括空闲区块，进而可以为存储系统420的地址映射提供存储空间。在下文中，将参见图8A和图8B描述有关数据移动的更多细节。

[0055] 图8A示意性示出了根据本公开的一个示例性实现的用于在多个存储设备之间移动数据的框图800A。如图8A中的箭头820所示，可以将存储设备260中的区块810中的数据移动至存储设备410中的区块812。图8B示意性示出了已经执行图8A所示的数据移动操作之后的各个存储设备的状态的图示800B。如图8B所示，此时存储设备260中的区块810已经变为空闲区块，而存储设备410中的区块812变为被使用的区块。可以不断地执行数据移动操作，直到资源池270中的已分配部分以均匀的方式分布在各个存储设备中。

[0056] 利用上述示例性实现，原有存储设备210、220、230、240、250、…、260中的数据被移动至新增的存储设备410，一方面可以平衡资源池270中的各个存储设备的工作负载；另一方面，还可以为存储系统420的地址映射提供更多的存储空间。

[0057] 根据本公开的一个示例性实现，存储系统420是基于独立磁盘冗余阵列的存储系统，存储系统包括多个切片(slice)。图9A示意性示出了根据本公开的一个示例性实现的存储系统的结构900A的示意图。如图9A所示，存储系统420可以包括多个切片，例如切片910、912、916。在此切片是指存储系统420中的比区块具有更小粒度的存储单位。根据存储系统420的配置，可以将切片的大小设置为预定的数值。例如，可以将一个区块划分为多个切片，并且可以以切片为单位来响应于来自存储系统420的用户的请求。以此方式，可以以更为精细的粒度来向用户分配存储空间，进而提高存储系统420的使用效率。

[0058] 根据本公开的一个示例性实现，如果接收到针对多个切片中的切片的切片访问请

求，可以基于存储系统的切片分配表确定切片的地址范围。继而，可以基于切片的地址范围确定虚拟地址范围660，并且从分配的存储空间中选择存储空间以用于切片。在此实现中，被分配的存储空间可以具有较大的范围，例如被分配的存储空间为与区块相关联的大小。此时，可以从被分配的存储空间中选择适合于切片大小的存储空间，并且将选择的存储空间分配给该切片。

[0059] 利用上述示例性实现，在接收到关于一个切片的切片访问请求后，即可为该切片分配存储空间。该切片访问请求将会触发针对大于针对切片大小的虚拟地址范围的存储空间的请求，并且将会向存储系统420分配大于切片大小的存储空间。此时，还可以将分配的存储空间中的未被该切片使用的部分分配给存储系统420中的其他切片。以此方式，可以加速存储系统420的地址映射422的创建，进而降低用户等待时间并提高存储系统420的性能。在下文中，将参见图9A和图9B描述有关切片的更多操作细节。

[0060] 继续参见图9A，存储系统420还可以包括与多个切片910、912、916相关联的位图920。在此位图920可以包括多个比特位，并且比特位的数量可以等于存储系统420中包括的切片的数量。位图920中的相应位指示多个切片中的相应切片中的数据是否为零。在下文中，将参见图9B描述有关位图920的更多细节。

[0061] 图9B示意性示出了根据本公开的一个示例性实现的存储系统中的切片与位图之间的关系的示意图900B。如图9B所示，位图920中的位922可以指示切片910中的数据是否为零；位图920中的位924可以指示切片912中的数据是否为零；位图920中的位926可以指示切片916中的数据是否为零。根据本公开的一个示例性实现，可以定义以数值“1”来表示相对应的切片中的数据为零。根据本公开的一个示例性实现，还可以定义以数值“0”来表示相对应的切片中的数据为零。

[0062] 根据本公开的一个示例性实现，为了加快对于存储系统420的访问速度，还可以首先访问位图920以确定待访问的切片中的数据是否为零。如果位图920指示切片中的数据为零，则可以不再访问切片中的数值。因而，还可能会存在位图920所对应的虚拟地址范围不可访问的情况。此时，需要首先针对位图920分配存储空间。根据本公开的一个示例性实现，可以确定与切片的地址范围相关联的位图的地址。进而，可以采用上文描述的方法，来将资源池270中的存储设备中的空闲空间分配用于位图920。

[0063] 根据本公开的一个示例性实现，由于所分配的存储空间的大小可能远大于位图920的大小，因而可以在分配的存储空间中选择位图空间以用于存储位图920。

[0064] 初始时，由于存储系统420为空中并不包括用户写入的数据，此时可以将位图920中的位设置为指示切片中包括的数据为零。在接收到针对目标切片的写请求后，可以更新与目标切片相对应的位的数值。根据本公开的一个示例性实现，如果确定用户访问的为写操作，可以向切片写入由写操作指定的目标数据，并且将位图中的与切片相关联的位设置为指示切片中的数据为非零。利用上述示例性实现，当已经向切片写入目标数据后，则可以将被写入数据的切片所对应的位设置为指示切片中的数据是非零数据。

[0065] 根据本公开的一个示例性实现，存储设备410是在资源池270的扩展期间被插入资源池270的新存储设备，并且新的存储设备410中的至少一部分存储空间的物理地址尚未与存储系统建立地址映射关系。尽管此时已经向资源池270中添加了新的存储设备410，此时该存储设备410中的空间对于存储系统420并不可用。利用上述示例性实现，可以利用资源

池270中的空闲空间，为存储系统420中的一部分创建地址映射422。以此方式，可以降低用户等待存储系统420完成地址映射操作的时间，进而提高存储系统420的性能。

[0066] 根据本公开的一个示例性实现，可以按照存储系统420中的多个切片的顺序，针对多个切片中的每个切片生成切片访问请求。利用上述示例性实现，利用上述示例性实现，一方面可以按照访问请求，来优先地为所访问的切片分配存储空间；另一方面，可以逐一地为每个切片分配存储空间，进而为存储系统420中的全部切片分配空间。

[0067] 在上文中已经参见图4至图9B详细描述了根据本公开的方法的示例，在下文中将参见图10详细描述相应的设备的实现。图10示意性示出了根据本公开的一个示例性实现的用于管理存储系统的设备1000的框图。存储系统与来自资源池中的多个存储设备中的至少一部分存储设备相关联，多个存储设备中的存储设备中的至少一部分存储空间对于存储系统不可访问。具体地，该设备1000包括：标识模块1010，配置用于将存储设备中的至少一部分存储空间内的区块标识为空闲区块；选择模块1020，配置用于响应于确定访问请求指定的存储系统中的虚拟地址范围不可访问，从多个存储设备中选择分别具有空闲区块的一组存储设备；映射模块1030，配置用于将虚拟地址范围映射至选择的一组存储设备中的空闲区块的物理地址范围；以及分配模块1040，配置用于向存储系统分配物理地址范围指定的存储空间。在此的设备1000可以配置用于执行上文描述的方法500中的各个步骤，在此不再赘述。

[0068] 图11示意性示出了根据本公开的一个示例性实现的用于管理存储系统的设备1100的框图。如图所示，设备1100包括中央处理单元(CPU)1101，其可以根据存储在只读存储器(ROM)1102中的计算机程序指令或者从存储单元1108加载到随机访问存储器(RAM)1103中的计算机程序指令，来执行各种适当的动作和处理。在RAM1103中，还可存储设备1100操作所需的各种程序和数据。CPU 1101、ROM 1102以及RAM 1103通过总线1104彼此相连。输入/输出(I/O)接口1105也连接至总线1104。

[0069] 设备1100中的多个部件连接至I/O接口1105，包括：输入单元1106，例如键盘、鼠标等；输出单元1107，例如各种类型的显示器、扬声器等；存储单元1108，例如磁盘、光盘等；以及通信单元1109，例如网卡、调制解调器、无线通信收发机等。通信单元1109允许设备1100通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0070] 上文所描述的各个过程和处理，例如方法500，可由处理单元1101执行。例如，在一些实现中，方法500可被实现为计算机软件程序，其被有形地包含于机器可读介质，例如存储单元1108。在一些实现中，计算机程序的部分或者全部可以经由ROM 1102和/或通信单元1109而被载入和/或安装到设备1100上。当计算机程序被加载到RAM 1103并由CPU 1101执行时，可以执行上文描述的方法500的一个或多个步骤。备选地，在其他实现中，CPU 1101也可以以其他任何适当的方式被配置以实现上述过程/方法。

[0071] 根据本公开的一个示例性实现，提供了一种用于管理存储系统的设备，包括：至少一个处理器；易失性存储器；以及与至少一个处理器耦合的存储器，存储器具有存储于其中的指令，指令在被至少一个处理器执行时使得设备执行用于管理存储系统的动作。存储系统与来自资源池中的多个存储设备中的至少一部分存储设备相关联，多个存储设备中的存储设备中的至少一部分存储空间对于存储系统不可访问，该动作包括：将存储设备中的至少一部分存储空间内的区块标识为空闲区块；响应于确定访问请求指定的存储系统中的虚

拟地址范围不可访问，从多个存储设备中选择分别具有空闲区块的一组存储设备；将虚拟地址范围映射至选择的一组存储设备中的空闲区块的物理地址范围；以及向存储系统分配物理地址范围指定的存储空间。

[0072] 根据本公开的一个示例性实现，存储系统是基于独立磁盘冗余阵列的存储系统，其中从多个存储设备中选择分别具有空闲区块的一组存储设备包括：基于独立磁盘冗余阵列的配置来确定一组存储设备的数量；以及基于数量来选择一组存储设备。

[0073] 根据本公开的一个示例性实现，基于数量来选择一组存储设备包括：从多个存储设备中确定包括空闲区块的存储设备的空闲设备数量；响应于确定空闲设备数量满足数量，从多个存储设备中选择包括空闲区块的存储设备以作为一组存储设备。

[0074] 根据本公开的一个示例性实现，基于数量来选择一组存储设备进一步包括：响应于确定空闲设备数量不满足数量，将存储设备以外第一存储设备中的第一区块中的数据移动至存储设备；将第一区块标识为空闲区块；以及增加空闲设备数量。

[0075] 根据本公开的一个示例性实现，该动作进一步包括：将存储设备以外的第一存储设备中的第一区块中的数据移动至存储设备；以及将第一区块标识为空闲区块。

[0076] 根据本公开的一个示例性实现，存储系统是基于独立磁盘冗余阵列的存储系统，存储系统包括多个切片，方法进一步包括：响应于接收到针对多个切片中的切片的切片访问请求，基于存储系统的切片分配表确定切片的地址范围；基于切片的地址范围确定虚拟地址范围；以及其中向存储系统分配物理地址范围指定的存储空间包括：从分配的存储空间中选择存储空间以用于切片。

[0077] 根据本公开的一个示例性实现，基于切片的地址范围确定虚拟地址范围包括：确定与切片的地址范围相关联的位图的地址，位图中的相应位指示多个切片中的相应切片中的数据是否为零；以及基于位图的地址来确定虚拟地址范围。

[0078] 根据本公开的一个示例性实现，该动作进一步包括：在分配的存储空间中选择位图空间以用于存储位图；以及响应于确定用户访问的为写操作，向切片写入由写操作指定的目标数据；以及将位图中的与切片相关联的位设置为指示切片中的数据为非零。

[0079] 根据本公开的一个示例性实现，存储设备是在资源池的扩展期间被插入资源池的新存储设备；以及新存储设备中的至少一部分存储空间的物理地址尚未与存储系统建立地址映射关系。

[0080] 根据本公开的一个示例性实现，该动作进一步包括：针对多个切片中的每个切片生成切片访问请求。

[0081] 根据本公开的一个示例性实现，提供了一种计算机程序产品，计算机程序产品被有形地存储在非瞬态计算机可读介质上并且包括机器可执行指令，机器可执行指令用于执行根据本公开的方法。

[0082] 根据本公开的一个示例性实现，提供了一种计算机可读介质。计算机可读介质上存储有机器可执行指令，当机器可执行指令在被至少一个处理器执行时，使得至少一个处理器实现根据本公开方法。

[0083] 本公开可以是方法、设备、系统和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质，其上载有用于执行本公开的各个方面的计算机可读程序指令。

[0084] 计算机可读存储介质可以是可以保持和存储由指令执行设备使用的指令的有形

设备。计算机可读存储介质例如可以是——但不限于——电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、静态随机存取存储器(SRAM)、便携式压缩盘只读存储器(CD-ROM)、数字多功能盘(DVD)、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身,诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波(例如,通过光纤电缆的光脉冲)、或者通过电线传输的电信号。

[0085] 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备,或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令,并转发该计算机可读程序指令,以供存储在各个计算/处理设备中的计算机可读存储介质中。

[0086] 用于执行本公开操作的计算机程序指令可以是汇编指令、指令集架构(ISA)指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码,编程语言包括面向对象的编程语言—诸如Smalltalk、C++等,以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络—包括局域网(LAN)或广域网(WAN)—连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。在一些实现中,通过利用计算机可读程序指令的状态信息来个性化定制电子电路,例如可编程逻辑电路、现场可编程门阵列(FPGA)或可编程逻辑阵列(PLA),该电子电路可以执行计算机可读程序指令,从而实现本公开的各个方面。

[0087] 这里参照根据本公开实现的方法、装置(系统)和计算机程序产品的流程图和/或框图描述了本公开的各个方面。应当理解,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合,都可以由计算机可读程序指令实现。

[0088] 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理单元,从而生产出一种机器,使得这些指令在通过计算机或其他可编程数据处理装置的处理单元执行时,产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中,这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作,从而,存储有指令的计算机可读介质则包括一个制造品,其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

[0089] 也可以把计算机可读程序指令加载到计算机、其他可编程数据处理装置、或其他设备上,使得在计算机、其他可编程数据处理装置或其他设备上执行一系列操作步骤,以产生计算机实现的过程,从而使得在计算机、其他可编程数据处理装置、或其他设备上执行的

指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

[0090] 附图中的流程图和框图显示了根据本公开的多个实现的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分，模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或动作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

[0091] 以上已经描述了本公开的各实现，上述说明是示例性的，并非穷尽性的，并且也不限于所公开的各实现。在不偏离所说明的各实现的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，旨在最好地解释各实现的原理、实际应用或对市场中的技术的改进，或者使本技术领域的其他普通技术人员能理解本文公开的各实现。

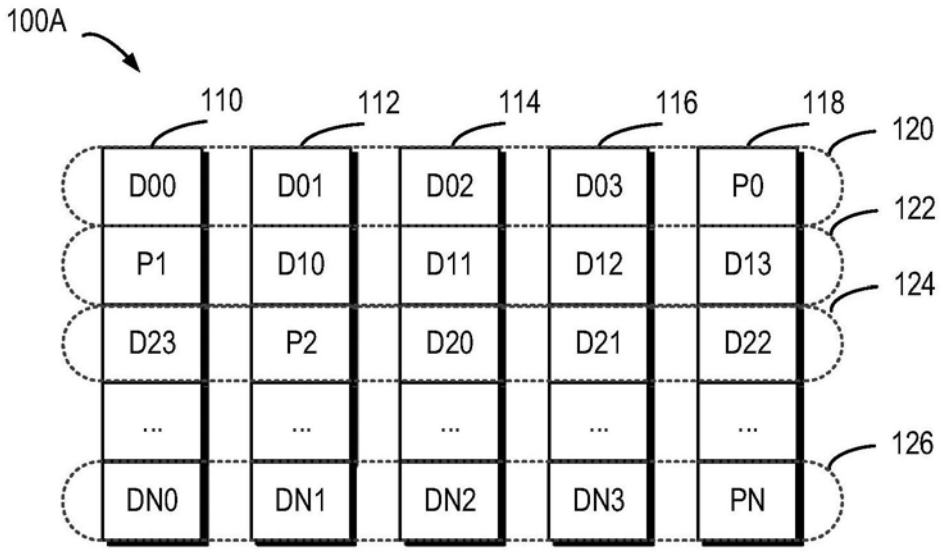

100A
110
112
114
116
118
120
122
124
126
D00
D01
D02
D03
P0
P1
D10
D11
D12
D13
D23
P2
D20
D21
D22
...
...
...
...
...
DN0
DN1
DN2
DN3
PN

图1A

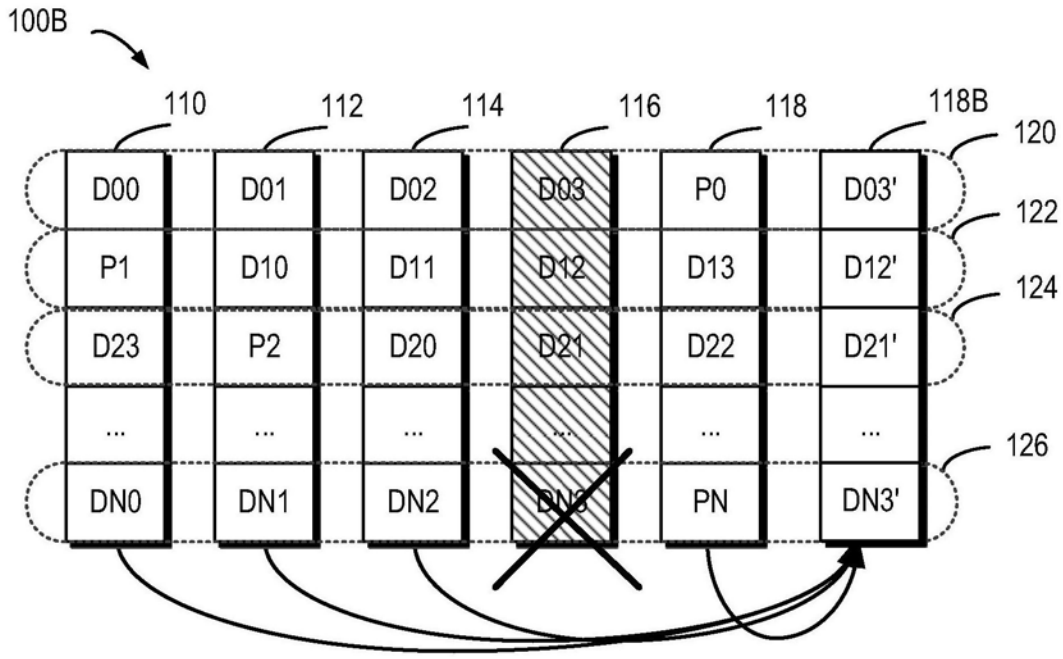

100B
110
112
114
116
118
118B
120
122
124
126
D00
D01
D02
D03
P0
D03'
P1
D10
D11
D12
D13
D12'
D23
P2
D20
D21
D22
D21'
...
...
...
...
...
...
DN0
DN1
DN2
DN3
PN
DN3'

图1B

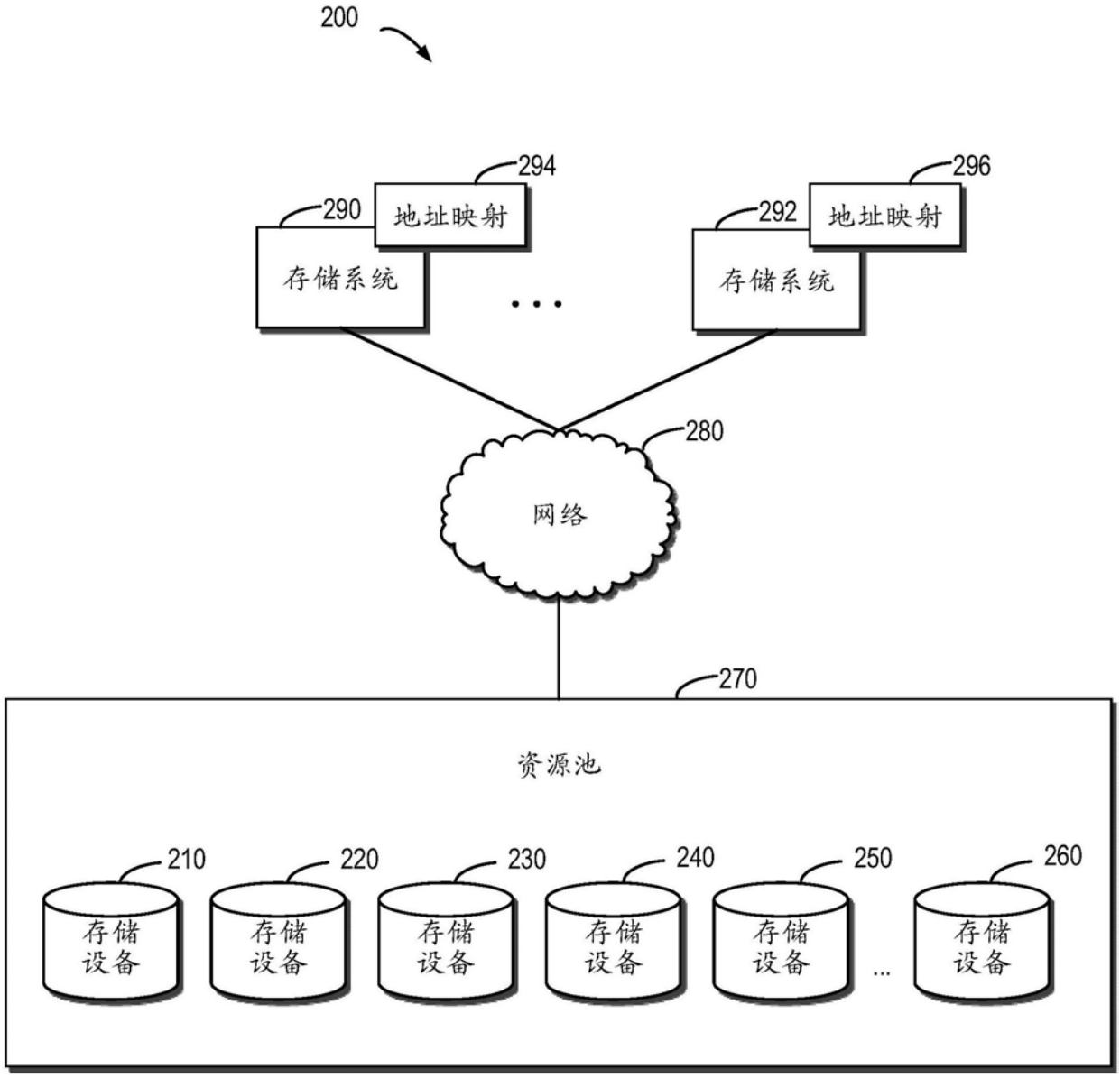

200
294
296
290
地址映射
292
地址映射
存储系统
存储系统
280
网络
270
资源池
210
220
230
240
250
260
存储设备
存储设备
存储设备
存储设备
存储设备
存储设备

图2

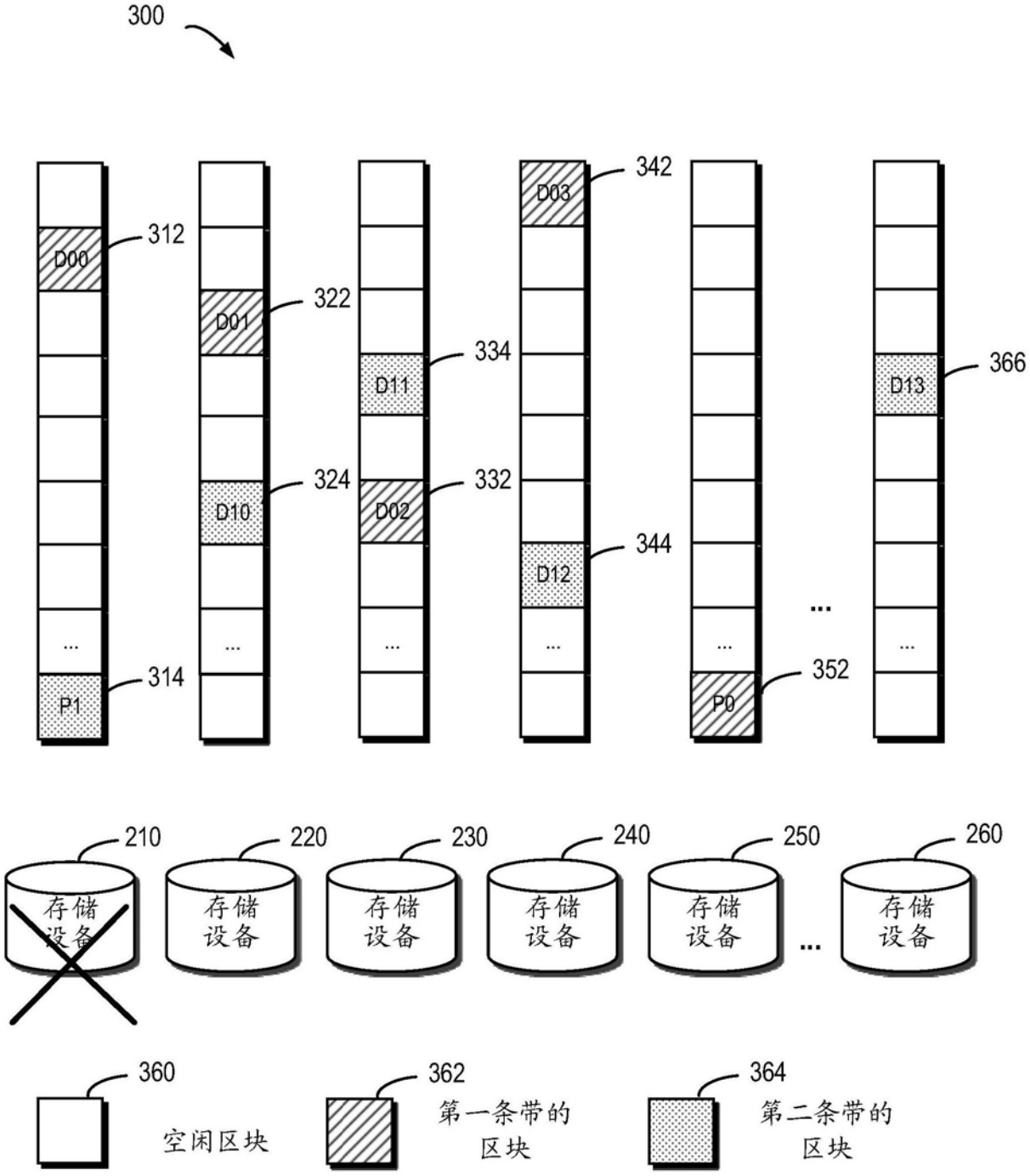

300
312
D00
322
D01
324
D10
314
P1
334
D11
332
D02
342
D03
344
D12
352
P0
366
D13
210
220
230
240
250
260
存储设备
360
空闲区块
362
第一条带的区块
364
第二条带的区块

图3

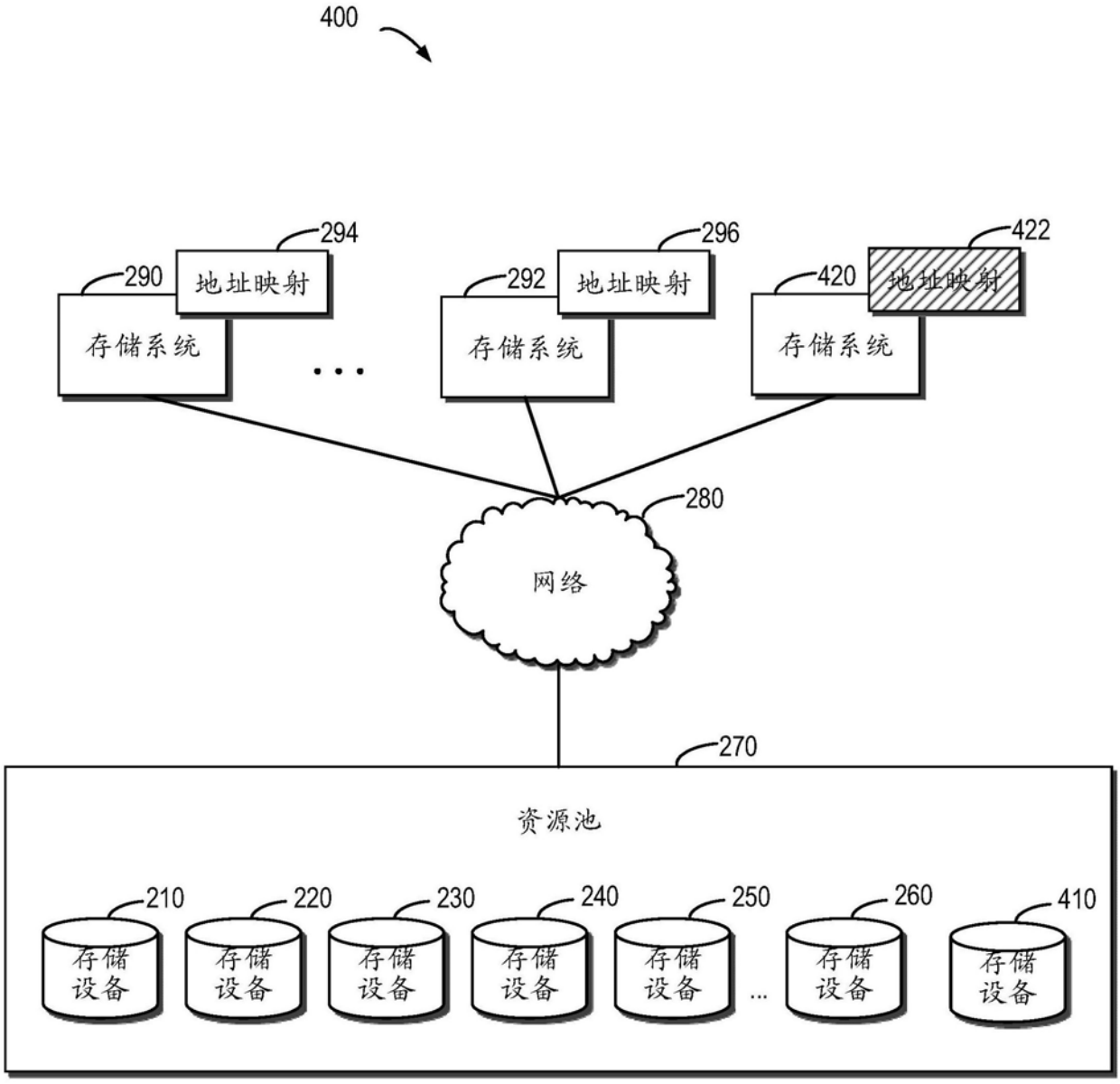

400
294
290
地址映射
存储系统
296
292
地址映射
存储系统
422
420
地址映射
存储系统
280
网络
270
资源池
210
220
230
240
250
260
410
存储设备
存储设备
存储设备
存储设备
存储设备
存储设备
存储设备

图4

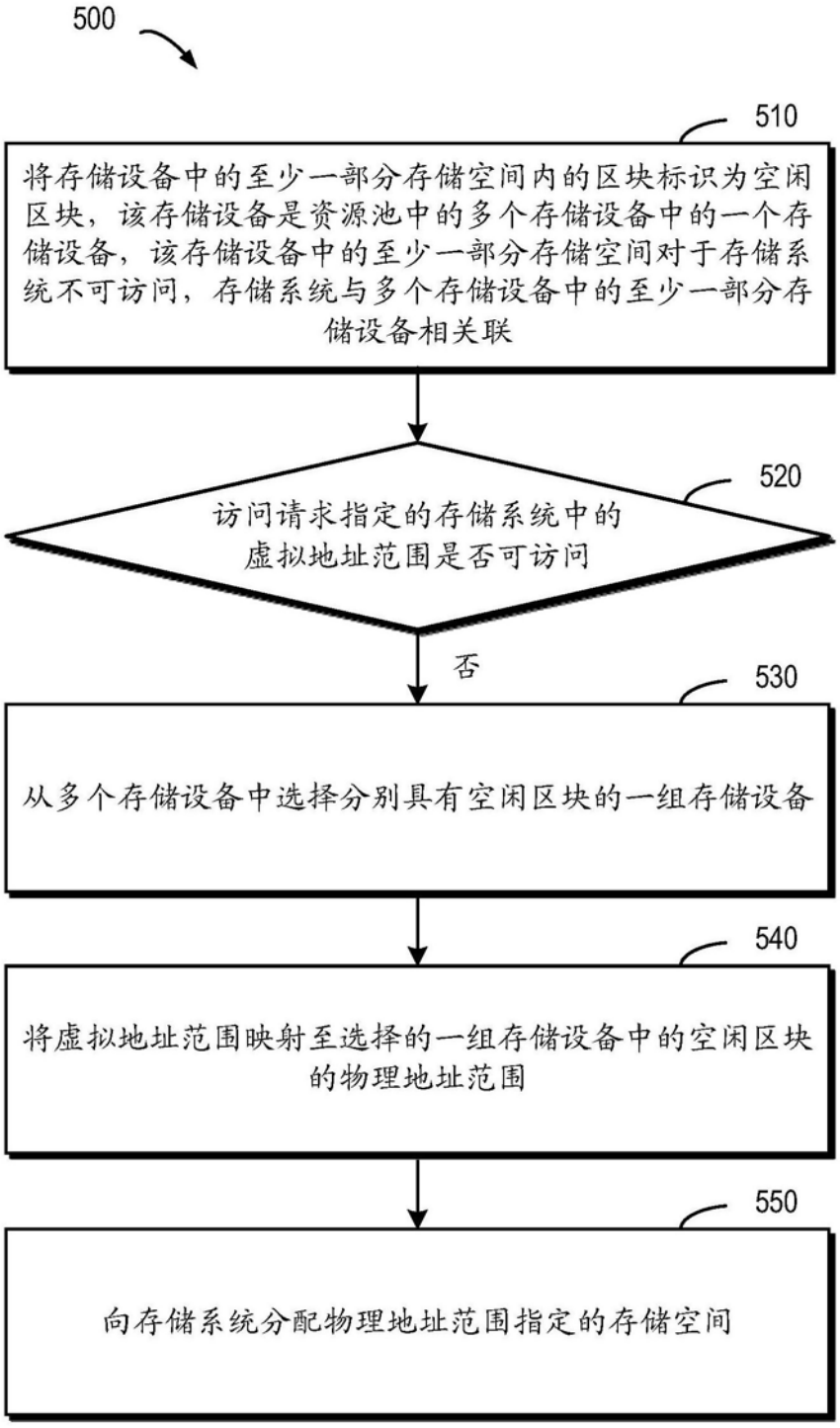

500
510
将存储设备中的至少一部分存储空间内的区块标识为空闲区块，该存储设备是资源池中的多个存储设备中的一个存储设备，该存储设备中的至少一部分存储空间对于存储系统不可访问，存储系统与多个存储设备中的至少一部分存储设备相关联
520
访问请求指定的存储系统中的虚拟地址范围是否可访问
否
530
从多个存储设备中选择分别具有空闲区块的一组存储设备
540
将虚拟地址范围映射至选择的一组存储设备中的空闲区块的物理地址范围
550
向存储系统分配物理地址范围指定的存储空间

图5

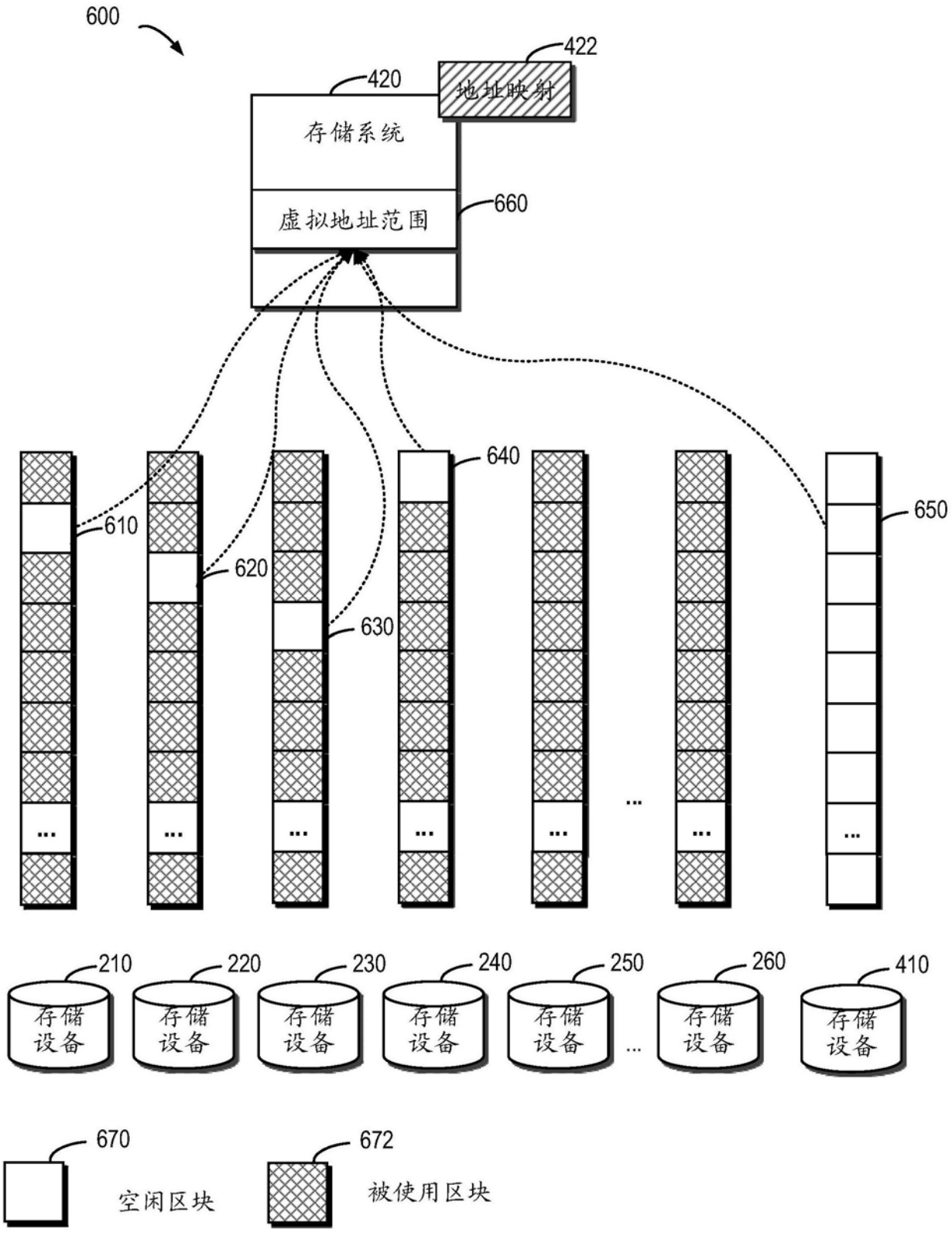

600
422
420
地址映射
存储系统
虚拟地址范围
660
610
620
630
640
650
210
220
230
240
250
260
410
存储设备
670
空闲区块
672
被使用区块

图6

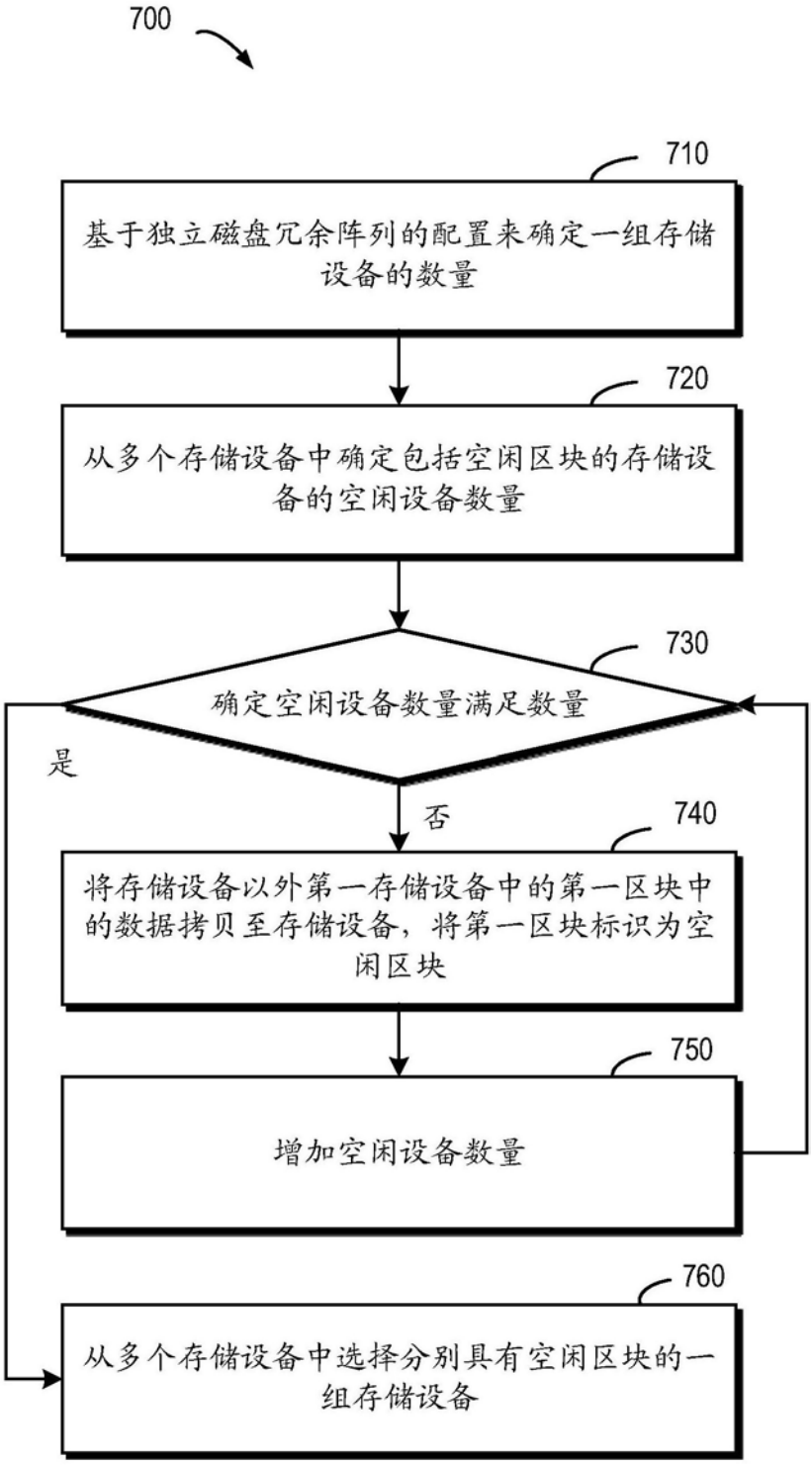

700
710
基于独立磁盘冗余阵列的配置来确定一组存储设备的数量
720
从多个存储设备中确定包括空闲区块的存储设备的空闲设备数量
730
确定空闲设备数量满足数量
是
否
740
将存储设备以外第一存储设备中的第一区块中的数据拷贝至存储设备，将第一区块标识为空闲区块
750
增加空闲设备数量
760
从多个存储设备中选择分别具有空闲区块的一组存储设备

图7

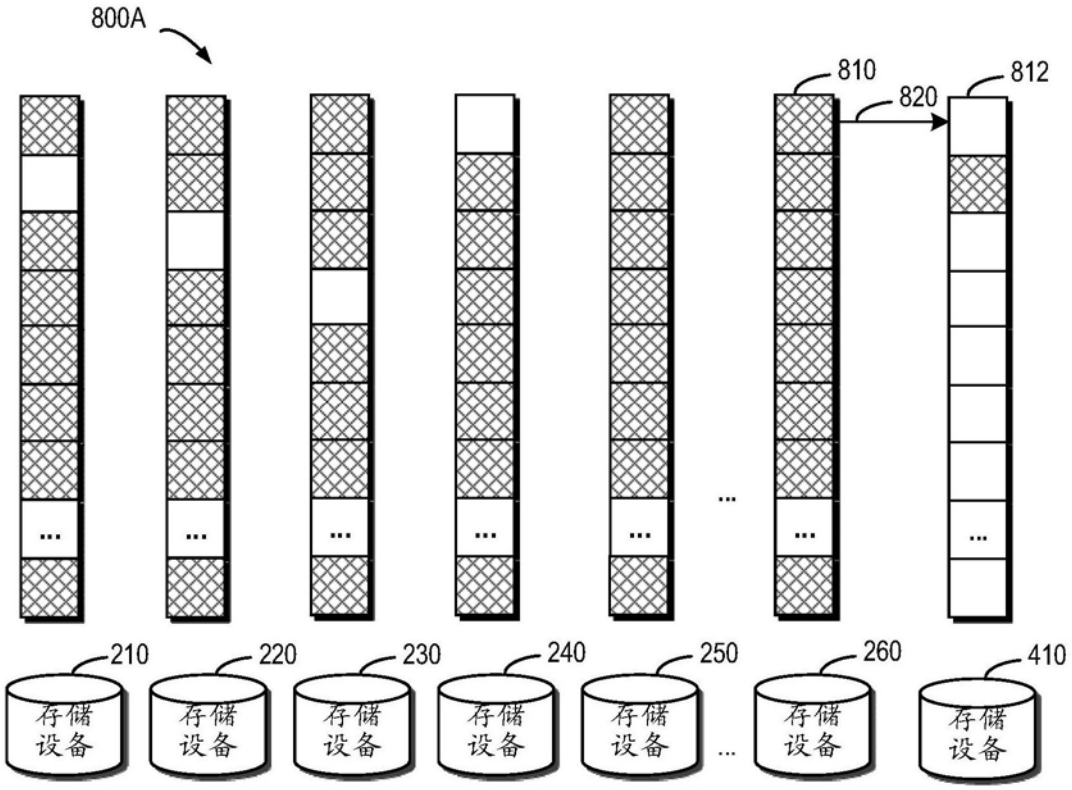

800A
810
820
812
210
220
230
240
250
260
410
存储设备

图8A

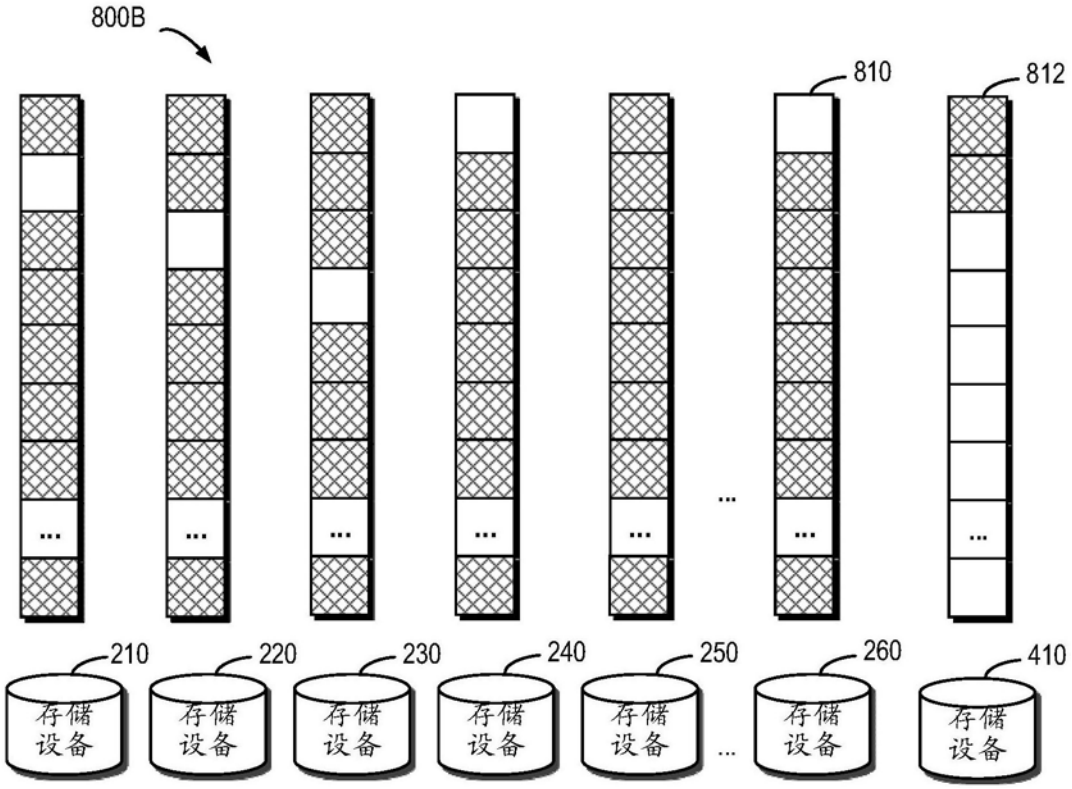

800B
810
812
210
220
230
240
250
260
410
存储设备

图8B

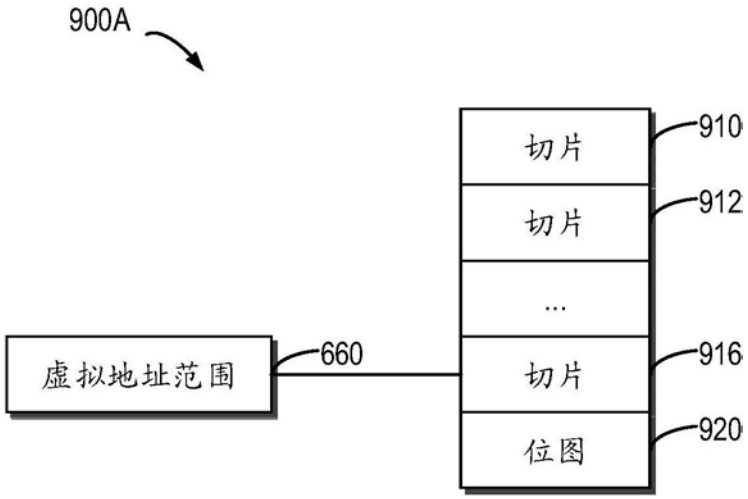

900A
切片
910
切片
912
...
虚拟地址范围
660
切片
916
位图
920

图9A

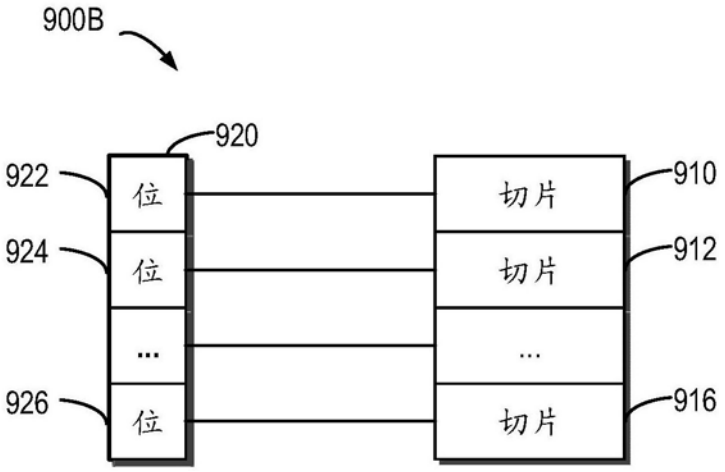

900B
920
922
位
924
位
...
926
位
切片
910
切片
912
...
切片
916

图9B

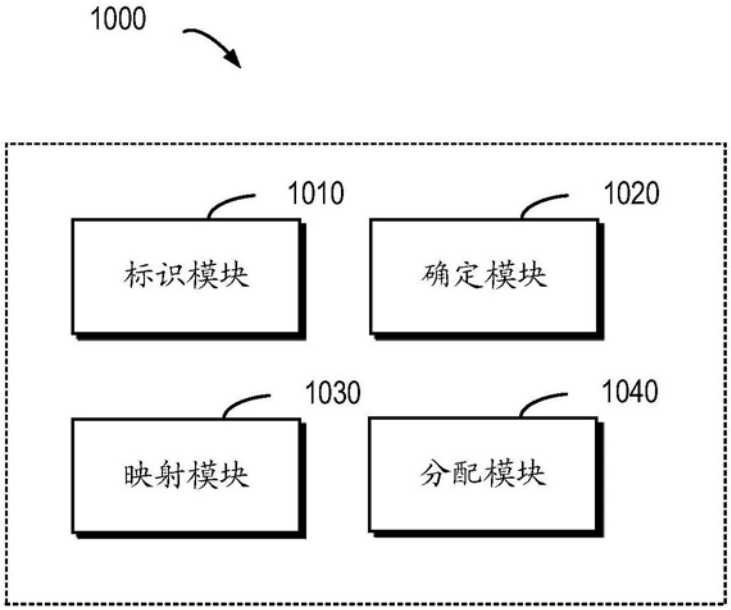

1000
1010
标识模块
1020
确定模块
1030
映射模块
1040
分配模块

图10

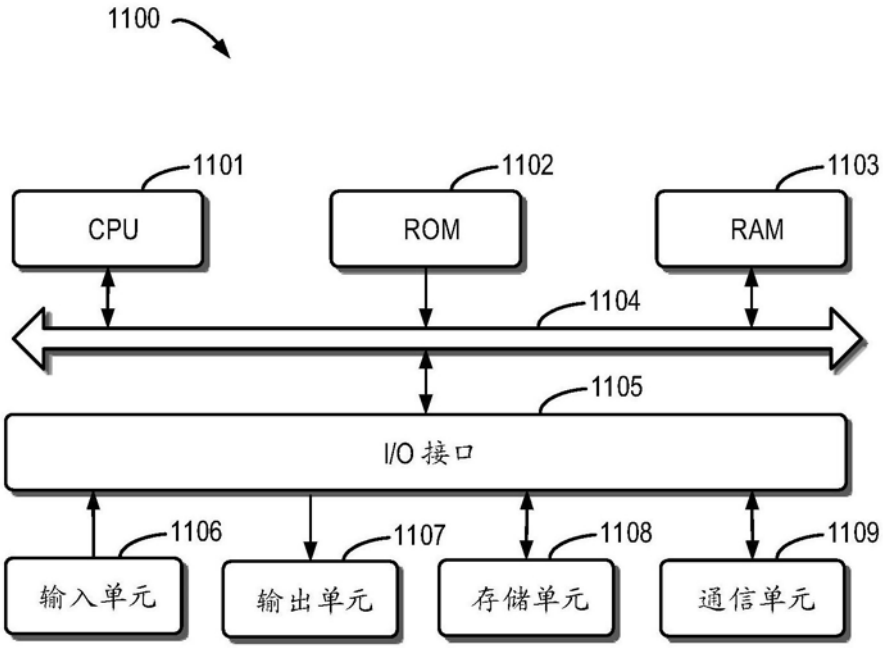

1100
1101
CPU
1102
ROM
1103
RAM
1104
1105
I/O 接口
1106
输入单元
1107
输出单元
1108
存储单元
1109
通信单元

图11